

USN []

NMAM INSTITUTE OF TECHNOLOGY, NITTE

(An Autonomous Institution affiliated to VTU, Belagavi)

VI Sem B.E. (CSE) Mid Semester Examinations - I, February 2017

14CS616 – MANAGING BIG DATA

Duration: 1 Hour

Max. Marks: 20

Note: Answer any One full question from each Unit.

Unit – I

- | | | |
|--|---------|--------|
| a) Describe two critical components of hadoop. | Marks 5 | BT* L4 |
| b) Explain the following | | |
| i) Fraud and big data | 5 | L2 |
| ii) Risk and big data | | |
| a) List variety of data sources in big data. | 3 | L3 |
| b) What is credit risk management? | 4 | L4 |
| c) Write three dimensions of big data. | 3 | L2 |

Unit – II

- | | | |
|--|---|----|
| a) What is impedance mismatch? What is key-value and document data models? | 5 | L3 |
| b) Write a note on graph databases. | 5 | L2 |
| a) Describe aggregate data model with a neat diagram. | 5 | L4 |
| b) List features of NoSQL.what are schemaless databases. | 5 | L2 |

Bloom's Taxonomy, L* Level

NMAM INSTITUTE OF TECHNOLOGY, NITTE
(An Autonomous Institution affiliated to
Karnataka State Board of Secondary Education)

(An Autonomous Institution affiliated to VTU, Belagavi)
B.E. (CSE) Mid Semester Exam.

VI Sem B.E. (CSE) Mid Semester Examinations - II, March 2015

14CS616 – MANAGING BIG DATA

Duration: 1 Hour

Max. Marks: 20

Note: Answer any **One** full question from each Unit

Unit - I

- Unit - I**

Ma

 - a) Describe Remote Procedure Call serialization Format.
 - b) Compare and contrast namenodes and datanodes. What are the benefits of compression?

 - a) Explain the design of Hadoop Distributed File System.
 - b) With a neat diagram explain internal structure of a sequence file with record compression and block compression.

Unit - II

- Ques. No. _____ Date _____

UNIT - 4

a) Describe shuffle and sort with a neat diagram. 7 L3
b) Explain task execution in MapReduce job run. 3 L2

a) With a neat Transition diagram explain how Apache Oozie is used for running workflows of dependent jobs. 5 L4
b) List five independent entities of MapReduce job run. 5 L2

Bloom's Taxonomy, L* Level

USN [REDACTED]
NMAM INSTITUTE OF TECHNOLOGY, NITTE
(An Autonomous Institution affiliated to VTU, Belagavi)

Sixth Semester B.E. (CSE) (Credit System) Degree Examinations
April – May 2017

14CS616 – MANAGING BIG DATA

Duration: 3 Hours

Max. Marks: 100

Note: Answer **Five full** questions choosing One full question from each Unit.

Unit – I

- | | | |
|--|-------|-----|
| a) With the help of the neat diagram explain Analytics spectrum. | Marks | BT* |
| b) Explain the fraud detection powered by near real time event processing with a neat diagram. | 10 | L*2 |
| a) Explain the Data in world of HealthCare. | 10 | L2 |
| b) Explain the Open source Technologies and Cloud & BigData | 10 | L2 |
| | 10 | L2 |

Unit – II

- | | | |
|---|----|----|
| a) List and explain common characteristics of NoSQL databases. | 6 | L1 |
| b) Define Aggregate and aggregate-ignorant. | 4 | L1 |
| c) Explain the example of Relations and Aggregates. | 10 | L2 |
| a) Discuss Schemaless Databases. | 4 | L2 |
| b) Define Sharding. How Sharding puts different data on separate nodes explain with neat diagram. | 8 | L2 |
| c) Explain with a neat diagram Master-Slave Replication and Peer-to-Peer Replication. | 8 | L2 |

Unit – III

- | | | |
|--|----|----|
| a) Explain MapReduce data flow with a single and multiple reduce task. | 12 | L2 |
| b) Explain the Design of HDFS. | 8 | L2 |
| a) Explain HDFS Concepts. | 12 | L2 |
| b) Define Serialization and Deserialization, explain RPC serialization format. | 8 | L2 |

Unit – IV

- | | | |
|--|----|----|
| a) Explain Writing a Unit Test with MRUnit with example. | 10 | L2 |
| b) Explain the MapReduce Job Running on a Cluster | 10 | L2 |
| a) Explain how the Hadoop runs a MapReduce Job. | 10 | L2 |
| b) Explain with a neat diagram Shuffle and Sort in MapReduce | 10 | L2 |

Unit – V

- | | | |
|---|----|----|
| a) Explain the Problems associated with Relational Database Systems. | 5 | L2 |
| b) Explain with a neat diagram the major "modules" of Hive and how they work with Hadoop. | 10 | L2 |
| c) Explain Primitive Data Types and Collection Data Types in Hive. | 5 | L2 |
| a) Explain Dynamic partitions properties in Hive. | 5 | L2 |
| b) Explain all Mathematical functions available in Hive. | 10 | L2 |
| c) Explain Join Optimizations in Hive. | 5 | L2 |

Bloom's Taxonomy, L* Level

USN [REDACTED]

NMAM INSTITUTE OF TECHNOLOGY, NITTE
(An Autonomous Institution affiliated to VTU, Belagavi)

Sixth Semester B.E. (CSE) (Credit System) Degree Examinations
Make up / Supplementary Examinations – July 2017

14CS616 – MANAGING BIG DATA

Duration: 3 Hours

Max. Marks: 100

Note: Answer **Five full** questions choosing **One full** question from **each Unit**.

Unit – I

- | | | |
|---|-------|-----|
| a) Describe Big Data. Differentiate structured and unstructured data. | Marks | BT* |
| b) What are the critical components of hadoop? | 10 | L*3 |
| | 10 | L3 |
| a) What are the applications of Big Data in medicine and health care? Discuss cloud and big data. | 10 | L4 |
| b) Explain the following with reference to Big Data: | | |
| i. Fraud management | 10 | L2 |
| ii. Risk management | | |

Unit – II

- | | | |
|--|----|----|
| a) Describe aggregate data model. What is sharding? Explain in detail. | 10 | L3 |
| b) Compare and contrast master slave replication and peer-peer replication. What are schemaless databases? | 10 | L4 |
| a) Discuss partitioning and combining in MapReduce. | 10 | L3 |
| b) Explain how MapReduce calculations are done. | 10 | L2 |

Unit – III

- | | | |
|--|----|----|
| a) Explain design of Hadoop Distributed File System (HDFS). | 10 | L3 |
| b) Briefly explain Serialization. | 10 | L4 |
| a) Describe MapReduce data flow with single and multiple reduce tasks. | 10 | L3 |
| b) Discuss compression. How compression is used in MapReduce. | 10 | L4 |

Unit – IV

- | | | |
|---|----|----|
| a) Briefly explain various sorting techniques. | 10 | L3 |
| b) Give the anatomy of MapReduce job run. | 10 | L4 |
| a) How unit tests are conducted with MRUnit? Explain. | 10 | L2 |
| b) Describe MapReduce workflows in detail. | 10 | L3 |

Unit – V

- | | | |
|--|----|----|
| a) Describe the data types and file formats of Hive. | 10 | L3 |
| b) List and explain the problems with relational database systems. | 10 | L2 |
| a) Discuss HiveQL data manipulation. | | |
| b) Briefly explain HiveQL data definition. | | |

Bloom's Taxonomy, L* Level
